BISS Biodiversity Information Science and Standards

Biodiversity Information Science and Standards 3: e37083
doi: 10.3897/biss.3.37083

OPEN ACCESS

Conference Abstract

Multi-context Knowledge Base using Calculated Descriptors from Xper3: the Archaeocyaths Knowledge Base example

Adeline Kerner[‡], Régine Vignes Lebbe[§]

‡ UMR 7207 – Centre de Recherches en Paléontologie - Paris CNRS – Sorbonne Université - MNHN, Paris, France
§ Sorbonne Université, MNHN, CNRS, EPHE, Université des Antilles, Institut Systématique Évolution Biodiversité, ISYEB, Paris, France

Corresponding author: Adeline Kerner (kerner@mnhn.fr)

Received: 11 Jun 2019 | Published: 13 Jun 2019

Citation: Kerner A, Vignes Lebbe R (2019) Multi-context Knowledge Base using Calculated Descriptors from Xper3: the Archaeocyaths Knowledge Base example. Biodiversity Information Science and Standards 3: e37083. https://doi.org/10.3897/biss.3.37083

Abstract

Natural sciences need to make assertions about characteristics of taxa. Traits and qualities or descriptors and states become increasingly crucial as a resource for identification adapted to both scientists and the public. Specialists, non-specialists, and the general public need different strategies for accessing the information. Creating a knowledge base is time-consuming and adapting this base to several needs seems to increase the required time substantially. Specialists think that an identification tool requires a complete overhaul when they want to change the target audience or the language... so they do not wish to get involved. How to minimize creation and update time? How to aggregate data only once in order to create a single knowledge base with different access levels?

Strategies for integrating different patterns of descriptors in a single knowledge base are useful to modulate descriptions without loss of information and with the certainty that everything is up to date in each context. This multi-context knowledge base, derived from a single trait dataset, can generate descriptions adapted to different contexts and users. To address this issue, we propose to use calculated descriptors in order to have on the same knowledge base, different versions of descriptors that are updated automatically when the reference trait is modified. Calculated descriptors are a distinctive feature of Xper3. These

© Kerner A, Vignes Lebbe R. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

descriptors are automatically computed from other descriptors by using logical operators (boolean operators).

Xper3 (http://www.xper3.com) is a web platform that manages descriptive data and provides interactive identification keys. Xper3 and its previous version, Xper2, have been used already for various taxonomic groups. We will focus on fossils in order to reveal how calculated descriptors in Xper3 knowledge bases can solve the multi-context problem. The main source of content is the archaeocyaths knowledge base (http://archaeocyatha.infosyslab.fr).

Archaeocyaths are the first animals to build reefs during the Cambrian. They are calcified sponges without spicules. The archaeocyaths knowledge base is an efficient resource for scientific studies and a useful tool for non-specialists, especially with the support of calculated descriptors. Correspondence between archaeocyath and sponge morphologies is not ready yet, but everything will be included into PORO (The Porifera Ontology, http://purl.obolibrary.org/obo/poro/releases/2014-03-06/) in the short term (an anatomy ontology about sponges).

In this knowledge base, calculated descriptors are used to:

1. create a consistent multilingual interactive identification key (French and English are available and Russian is in draft),
2. generate descriptors adapted to different level of expertise and
3. reword morphological descriptors (adapted for identification) into homologous characters (adapted for phylogeny).

Xper2 and Xper3 are compatible with TDWG's Structured Descriptive Data (SDD) format. Calculated descriptors do not exist in SDD format and so they are exported from Xper3 as categorical descriptors, therefore losing the origin of values. Calculated descriptors are powerful and we are interested in discussing these with SDD and Xper3 users in order to improve the user interface and develop new tools for the analysis of such descriptors.

Keywords

knowledge base, calculated descriptors, Xper3, Archaeocyatha, palaeontology, identification, logical operators

Presenting author

Adeline Kerner

Presented at

Biodiversity_Next 2019